

Результат анализа текста

*The Master and Margarita.txt

Количество
словоупотреблений

> 112 k

*112902

Количество
словоформ

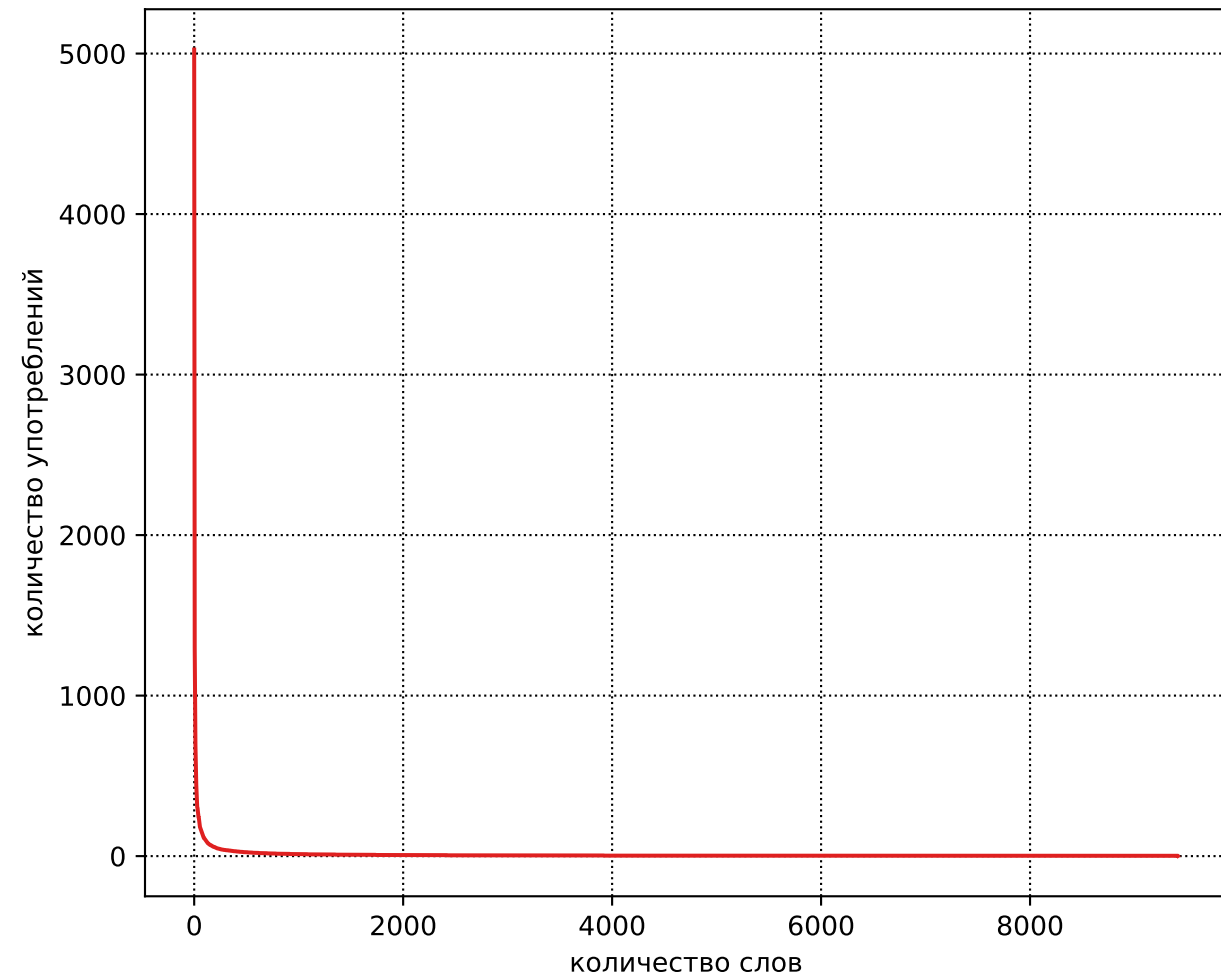
23913

Количество
лемм

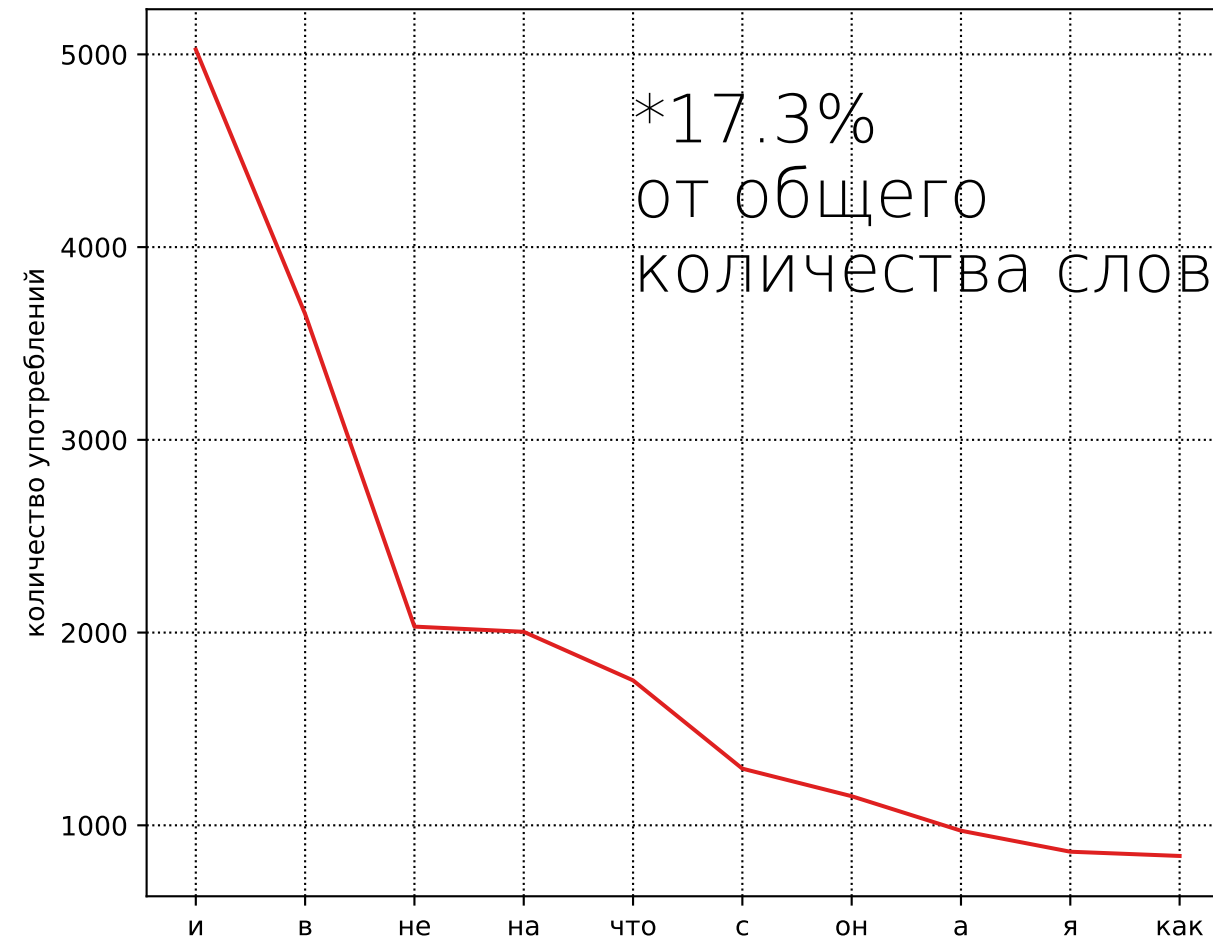
11234

*47.0% от числа
словоформ

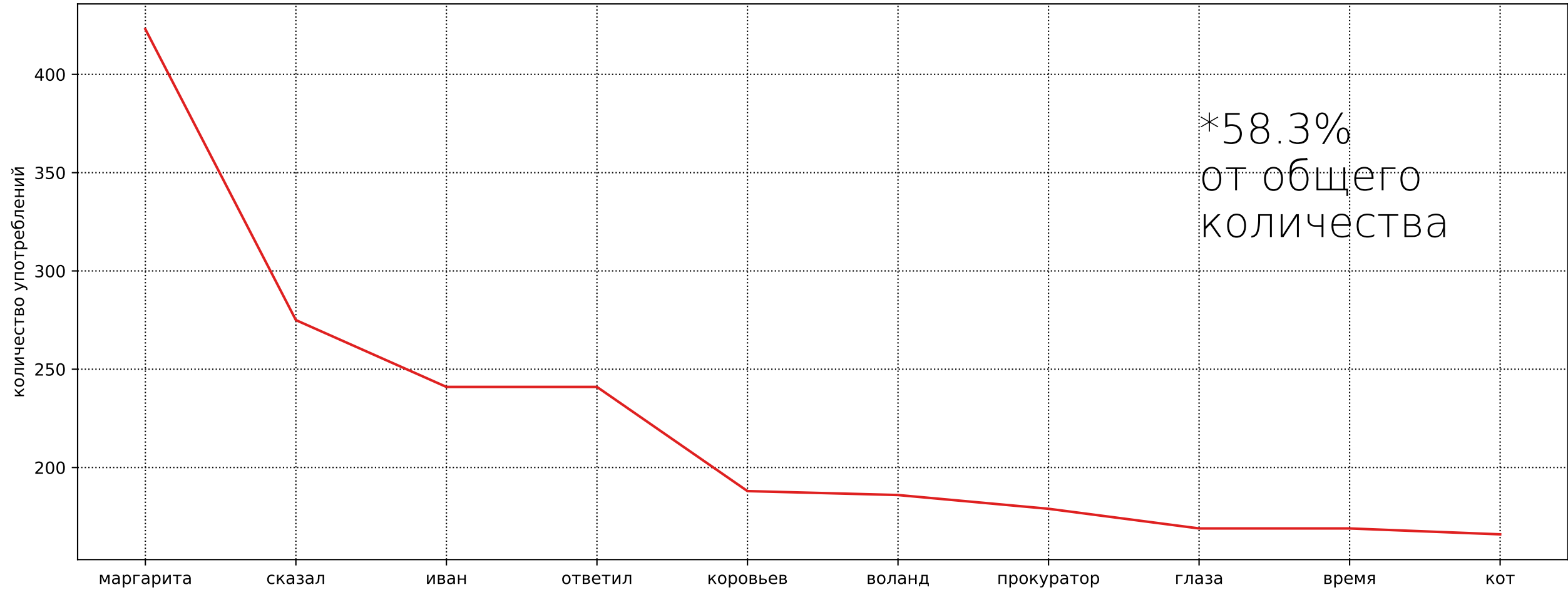
Характер распределения
слов в тексте (max, min)



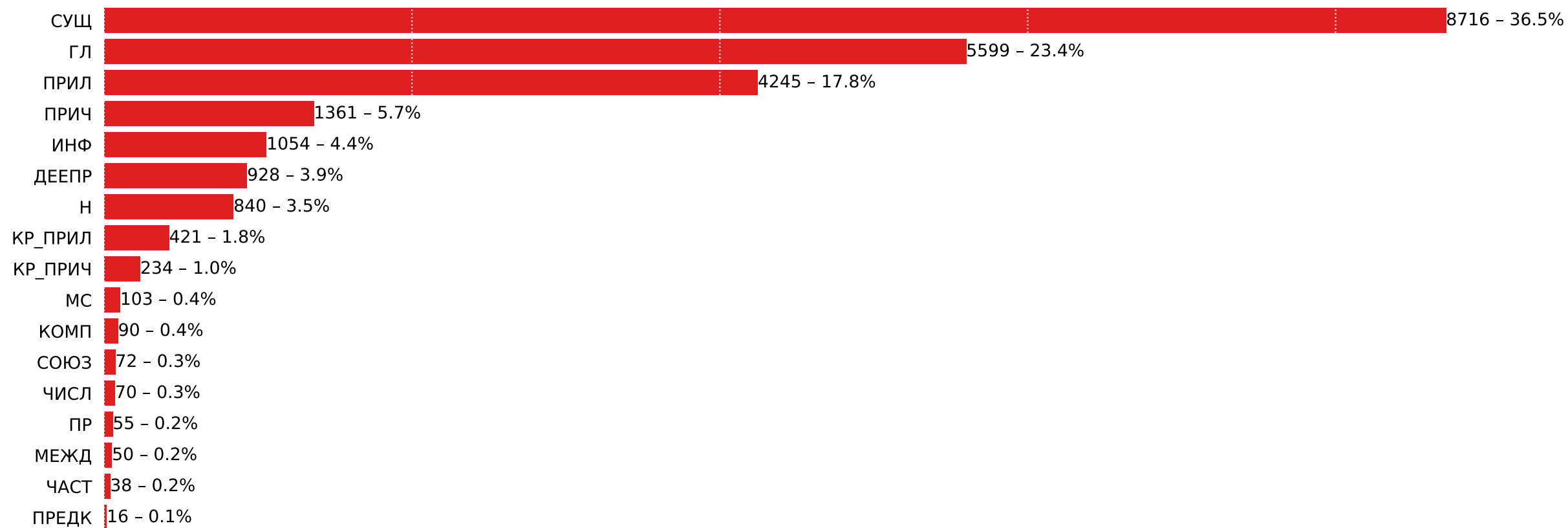
Распределение 10 наиболее частых
слов в тексте



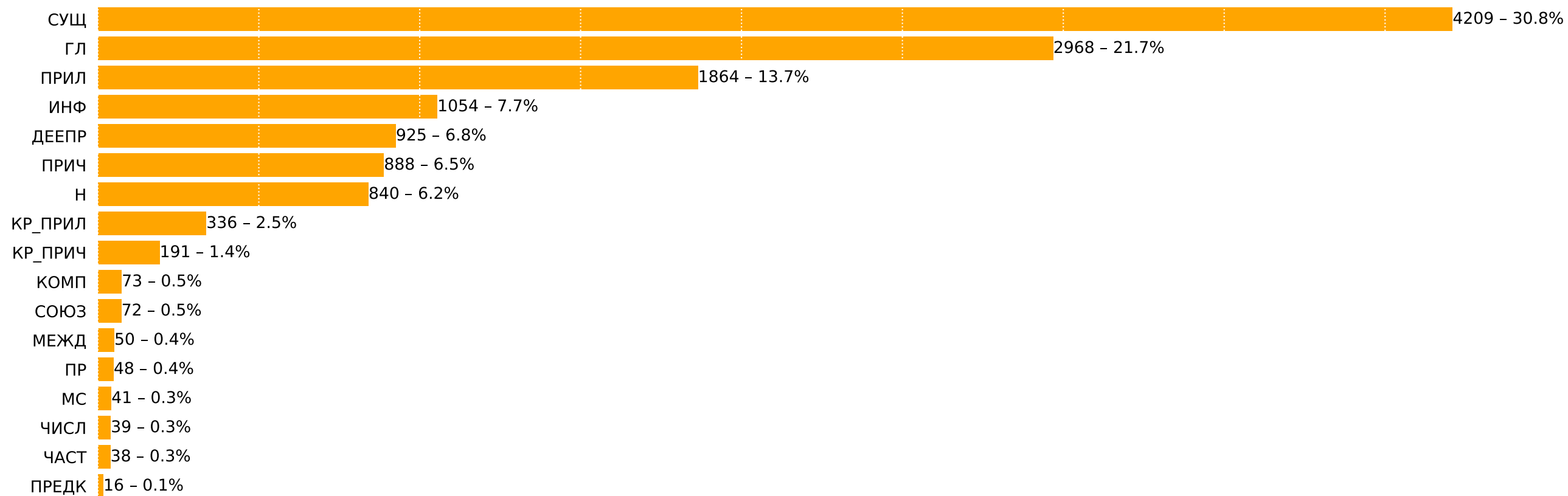
Распределение 10 наиболее частых слов в тексте (без стоп-слов)



Распределение словоформ по частям речи



Распределение лемм по частям речи



Коэффициент
лексического богатства текста

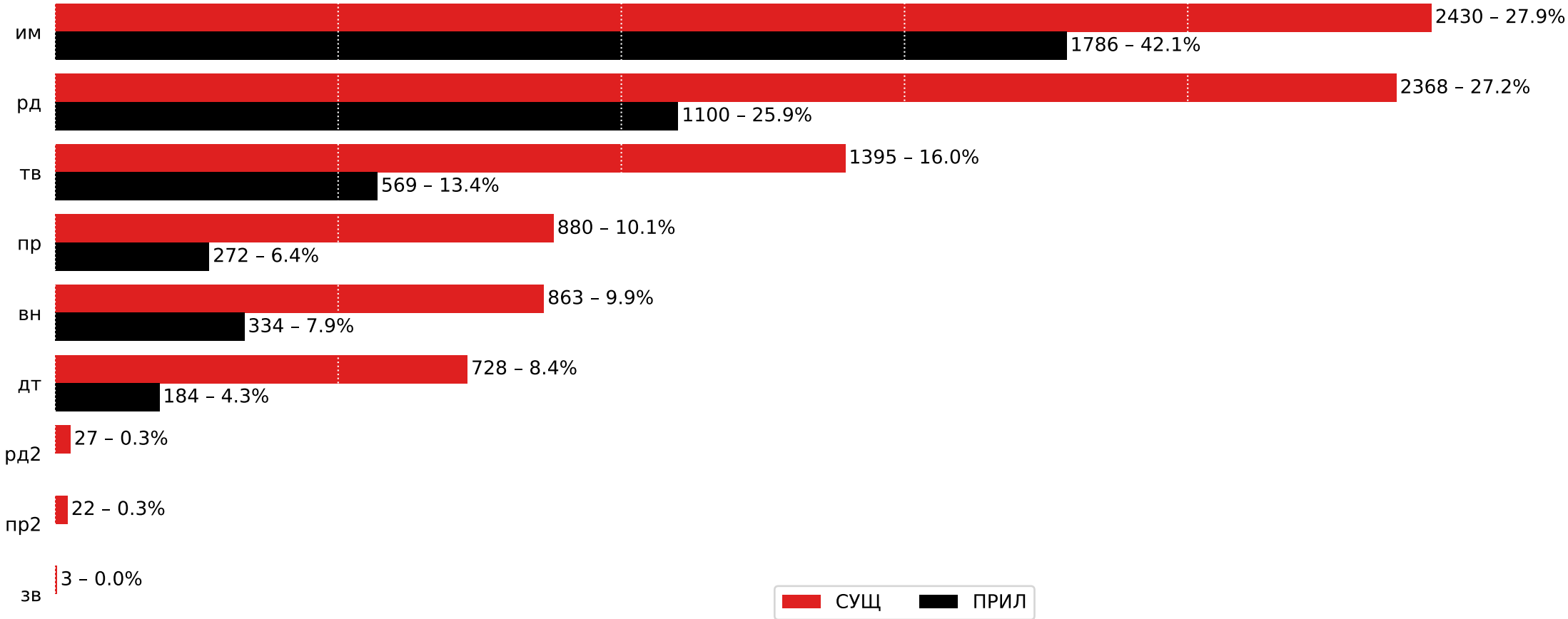
0.47

*отношение числа различных
лемм к общему числу словоформ

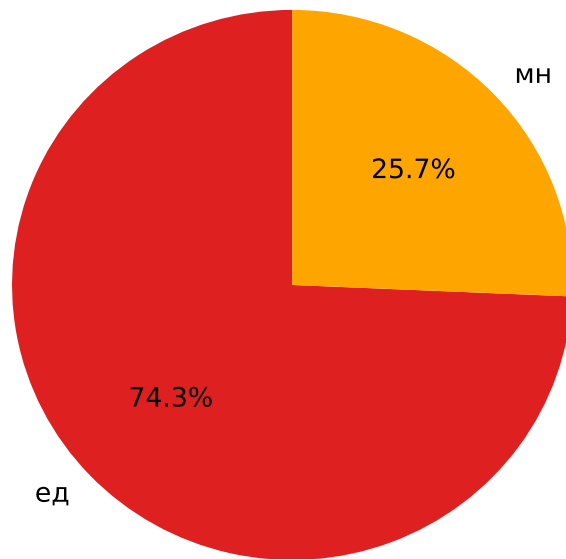
Процент
омонимичных словоформ

45.2%

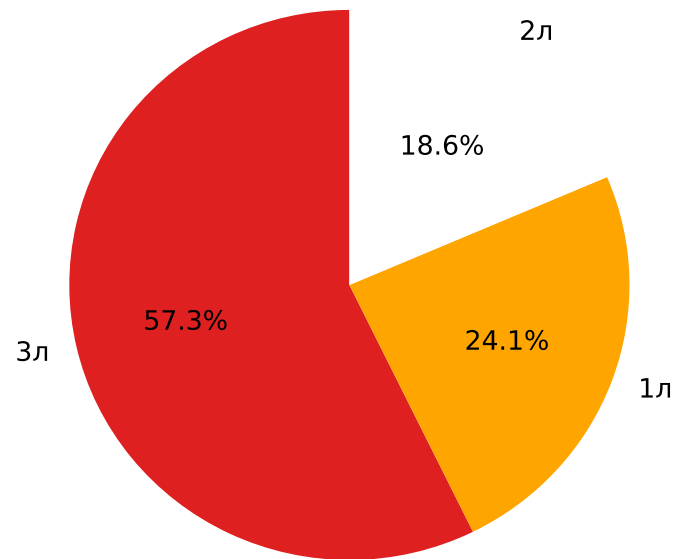
Распределение падежей
существительных и прилагательных



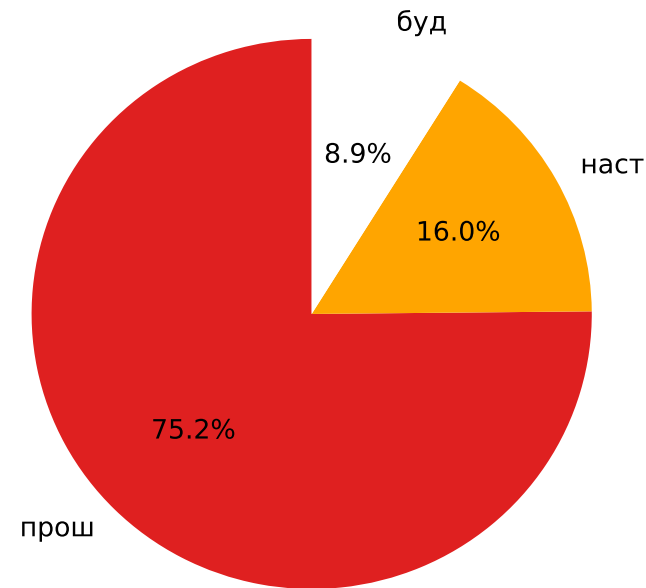
Распределение категории числа
у глаголов



Распределение категории лица
у глаголов



Распределение категории времени
у глаголов



Процент слов,
отсутствующих в словаре

0.8%

*таблица несловарных слов
содержится в файле
[./result/unknown_words_](#)

махинджаури | пятнажко | симпатичнейший |
парчевским | пороховниковой | искушающе |
ландшафтик | турма | решительнейшим |
фалернское | кефи | пятидесятилетняя |
выражается | себастиийских | соболезнающе |
ювелиршины | зеленохвостые | вашнер |
пятисвечий | коровьевские | лафитничек |
эх-хо-хо | дипломатически | желдыбиным |
хустова | вегнер | тссс |
желдыбин | истерически | хустовых |
оближ | недалечко | лиходеевской |
палисадничек | массолитским | кривоглазие |
четырехкомнатную | чесунче | препровожден |
груня | лукинишна | геркулановны |
пролежнева | зеленоспинных | помещеньице |
вареньке | спиридоновке | чердакчи |
беккеровский | ершалаимского | степиною |
намучавшись | аплодисмент | скабичевский |
<...>